Before the
Federal Communications Commission
Washington, D.C. 20554

In the Matter of	)	
	)	
Video Description: Implementation of the Twenty-First Century Communications and Video Accessibility Act of 2010	)))	MB Docket No. 11-43

Comments of Dolby Laboratories, Inc.

I. Introduction and Summary

Dolby Laboratories (Dolby) directs its comments to Paragraph 31 of the Federal Communications Commission's (FCC's or Commission's) *Notice of Proposed Rule Making* released on March 3, 2011, which seeks comment on whether the FCC should update its rules to incorporate the ATSC digital standard A/53 Part 5:2010 which lacks the second option of a receiver mix contained in the ATSC's A/53 Part 5: 2007 which is currently referred to in FCC rules.

The second option, requiring support in DTV receivers, was eliminated by the ATSC in 2010 because the receiver mix approach had never been implemented in AC-3 decoders, and there was no expectation that the approach would ever be implemented in AC-3 decoders in the foreseeable future. However, it should be noted that a receiver-mix approach has been implemented in products implementing the more advanced E-AC-3 technology that is documented in ATSC A/52, and A/53 Part 6. The broadcast mix mode described in A/53 Part 5:2010 does not provide the best delivery of video description and Dolby is concerned that the exclusive incorporation into the FCC's rules of broadcast-mix as described in A/53 Part 5:2010 could foreclose the opportunity to better serve the visually-impaired audience. Dolby therefore encourages the FCC to avoid drafting the rules in such a way as to prevent the subsequent evolution to a "receiver-mix" solution.

II. **Discussion**

Audio description has long been provided by broadcast services. Until recently, the only alternative has been a broadcast-mix approach whereby main and descriptive tracks are mixed in postproduction prior to transmission, and typically provided as a monophonic service.

Today, however, a more efficient alternative is available in both broadcasting and receiving equipment whereby a descriptive track is transmitted as a separate bitstream alongside the main soundtrack. Mixing then occurs within the *receiver*, allowing the visually impaired viewer to enjoy the multi-channel audio soundtrack as well as the video description. This receiver-mix approach makes it possible to deliver audio descriptions for the visually impaired while maintaining the highest possible bandwidth efficiency and the richest audio experience.

Attached as Appendix A to these Comments is a paper entitled "Audio Mixing Requirements in Next Generation Broadcast Receivers for Audio Description and Other Enhanced Features" presented at the 62nd NAB Annual Broadcast Engineers Conference (April 12-17, 2008) which further describes the advantages and requirement of a receiver-mix solution.

Strong demand for a receiver-mix solution from public broadcasters in Europe has resulted in its inclusion in the DVB specification (DVB (Annex E of ETSI TS 101 154 v 1.9.1); DVB (Annex E of ETSI TS 102 366 v 1.2.1); DVB (Annex D of ETSI EN 300

468 v 1.11.1)) and implementation in selected markets. Appendix B to these Comments sets out the European HD Specifications. The fact that there was demand for, and specification and actual implementation of, receiver-mix technology in Europe has resulted in this technology becoming available in consumer receiver products at reasonable cost. As many television products are built for the world market, this technology is also available for use in U.S. products, thus it is highly practical to plan to use this technology in the future to optimize both the bandwidth used for the VI service, and the quality of service provided to the visually-impaired audience.

III. **Conclusion**

A receiver-mix approach offers significant quality advantages and broadband efficiencies over the broadcaster-mix solution currently described in ATSC Part 5: 2010. Those advantages and efficiencies are the reasons a receiver-mix solution is presently implemented in Europe; and, in turn, the implementation of the receiver mix technology in receivers for the European market makes the technology a practical alternative for the U.S. market as well. Dolby recommends that the FCC not perpetuate the less efficient technology and inferior listener experience of ATSC Part 5 as it is currently applied to video-described content, but instead draft its rules to allow for the transition to this improved receiver-mix technology offering greater audio quality to multiple audiences.

Dolby would welcome the opportunity to confer with the FCC and other industry stakeholders to discuss the reasonable implementation of, and mechanisms for, a transition to a receiver-mix solution.

Respectfully submitted,

Craig C. Todd

Craig Todd
Senior Vice President; Chief Technology Officer
Dolby Laboratories, Inc.
100 Potrero Avenue
San Francisco, CA 94103
(415) 645-5000

April 28, 2011

Appendix A

NAB2008

Audio Mixing Requirements in Next Generation Broadcast Receivers for Audio Description and Other Enhanced Features

Roland Vlaicu

Dolby Laboratories
San Francisco, CA

ABSTRACT

The deployment of next generation broadcast platforms for HDTV and IPTV offers the opportunity to utilize advanced video and audio codecs, as permitted by recent revisions to ETSI specifications TS 101154 and TS 102005. In addition to bandwidth savings, new audio codecs for these applications will likely need to support additional features within their platform lifetimes, including 7.1-channel surround sound and improved provision for the visually impaired. Support for the latter feature is increasingly being mandated by broadcast regulators in Europe and elsewhere.

This paper specifically considers the requirements for audio mixing features within next generation broadcast audio codecs. Audio mixing enables enhanced services, such as audio description for the visually impaired or director commentaries, to be offered at efficient total data rates. This can be achieved by transmitting an additional bit rate-efficient channel of commentary, which is mixed dynamically with the main program audio within the audio decoder of a standard home set-top box.

It is concluded that mixing audio streams in the coded domain within a single standard audio decoder offers significant advantages over other approaches. These include simplicity of implementation, improved audio quality and connectivity to home audio equipment. The importance of implementing metadata control of the mixing feature is highlighted, and requirements for appropriate production and encoding tools are discussed.

1. INTRODUCTION

This paper discusses a new approach to offer support for audio description (AD) and other advanced services in broadcast receiver devices such as set-top boxes or TV sets through the introduction of a mixing component within the audio decoder.

Mixing solutions in broadcast receivers that support two-channel mixing have been available for quite some time now. The main and associated audio tracks are presented using one of the common broadcast audio codecs for stereo such as MPEG-2 Layer II and then mixed at the user's option upon decoding. Some broadcast standards such as DVB [2] use defined mixing metadata for this purpose in order to support codec-agnostic mixing in the receiver device. This approach is often referred to as receiver-mix. However, in these cases functionality is limited, usually requiring the user has to choose between the audio description services and multichannel surround sound. Although the syntax of the DVB specification describes how to pan an audio description source within a 5.1 surround signal, it has not yet been implemented.

Next-generation broadcast specifications around the world focus on two codecs to distribute multichannel surround sound: Enhanced AC-3 and aacPlus. Both codecs support surround sound configurations of 5.1 and more at very efficient bit rates. While aacPlus can make use of the mixing metadata defined in [6], Enhanced AC-3 comes with its own set of mixing metadata that is contained in the associated audio stream. This brings many advantages, including the ability to include the associated audio stream as an independent substream with the main audio service, all signaled in the transport stream under the same Audio-PID.

2. AUDIO DESCRIPTION AS A STANDARD REQUIREMENT FOR NEXT-GENERATION BROADCAST RECEIVERS

In many countries, broadcast services for visually impaired audiences are becoming mandatory, and are provided mostly through an associated audio stream containing both the main audio mix and an audio description soundtrack. Preparation of the AD track is time-consuming in post production, and requires approximately the same bandwidth on the transmission path as the main audio stream. Enhanced AC-3 offers a more efficient alternative.

By integrating a mixing module with the audio decoder in a broadcast receiver device such as a set-top box or television set, Enhanced AC-3 enables bit-stream mixing in the discrete cosine transform (DCT) domain. This allows mixing an associated audio stream of up to 5.1 channels, as well system sounds and interactive audio of up to 5.1 channels, with the main audio service implemented on the audio decoder level. User level adjustment parameters enable the user to control the balance between main and associated audio in the overall output.

Such a mixing module could enable a variety of new features such as audio description via text-to-speech rendering, which many European public broadcasters employ for director's commentary with video-on-demand services. It could also enable the playback or generation of interactive sounds on the set-top box, with the resulting audio signal provided via both two-channel outputs and digital IEC 61937 interfaces (S/PDIF or HDMI). It would also assist broadcasters such as SVT in Sweden who are looking into offering an associated audio track derived from a text-to-speech engine, which could allow a dramatic increase in programming with AD services.

3. RECEIVER MIX VS. BROADCAST MIX

Audio description has long been provided by broadcast services. The PAL system, for example, makes it possible for the second FM sound sub-carrier, otherwise used for the right channel in stereo transmissions, to carry an alternative version of the main soundtrack carried on the first FM sub-carrier. While this "broadcast-mix" approach limits the transmission to mono main and associated audio soundtracks, it is an efficient way to provide a basic service for the visually impaired. Figures 1 shows the symbol used by European broadcasters to mark transmissions offering Audio description.

Figure 1: Symbol used by European broadcasters for transmissions offering audio description (AD)

This broadcast-mix approach, whereby the main and commentary tracks are mixed in postproduction prior to transmission, was carried over into digital television DTV systems because early receivers offered only basic functionality, and the sources for digital transmissions were often still stored on analog media. It is also similar to that taken by the NTSC system where a secondary audio program (SAP) carries a pre-mixed version of main and audio description tracks as a mono signal. However, due to the affiliate broadcast structure in the US, at times technical impediments interfere with passing through a SAP signal on cable and satellite systems, making the service less accessible to viewers.

In Germany, public broadcasters offering a large portion of their programming with audio description services are facing the challenge of continuing to distribute their AD service "analog-style", i.e. leaving main audio on the left channel and AD audio on the right channel of what is basically the main broadcast audio track carried as a stereo signal but signalled as "dual-mono". This requires broadcast receivers to pass only the left channel or right channel to their audio outputs. However, since this is not a requirement within DVB, which mandates that different audio services be transmitted and signaled using different audio-PID, receiving and recording these services becomes a major challenge.

To this date, transmitting audio description services in the form of a broadcast mix is standard practice in many parts of the world. Efforts to offer a more efficient way to provide services for the visually impaired in the form of a "receiver-mix" approach have cumulated in the creation of a specification that is now part of DVB.

Annex E of [2] describes a basic approach to transmit an isolated mono signal containing the description track alongside the main soundtrack as a separate bitstream. A suitable broadcast receiver can then perform the mixing of the main soundtrack with the commentary track in the device itself, hence the term receiver mix.

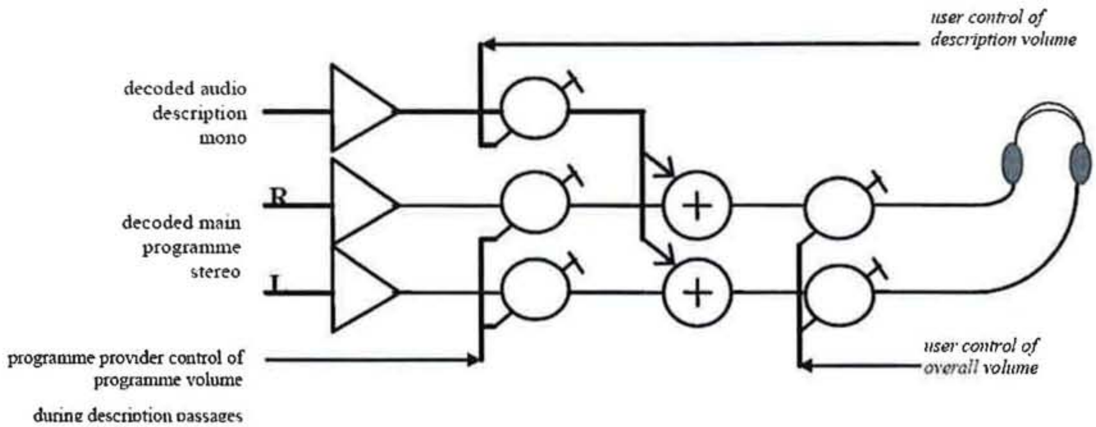

Figure 2: 2-channel Audio description as defined for DVB receivers in [2]

Figure 2 shows a diagram of the DVB approach. The DVB specification limits the associate audio stream to mono. While this approach is very valid to offer basic receiver-mix functionality to offer audio description services, it does not take into account other features that the mixer component in a broadcast receiver could support, such commentaries by several voices, or multichannel system sounds that could be overlaid onto a multichannel main soundtrack. The design of an enhanced mixing engine that supports these extended features is discussed in the next chapter.

There is strong demand for this feature from public broadcasters in Europe, and it can add another differentiating feature to broadcast devices that will help to drive the success of next-generation digital television services around the world.

4. PROPOSAL OF A NEW MIXING ARCHITECTURE

This chapter describes in detail a decoder-mixer-converter (DMC) for use in broadcast applications such as set-top boxes. The mixing of substreams with the main broadcast audio enables services such as audio description for the visually impaired or director's commentary, while at the same time maintaining the highest possible bandwidth efficiency.

Substream mixing negates the need for broadcasters to transmit several final mixes to support multiple audience preferences. It is necessary only to add low-bandwidth substreams containing the additional information along with mixing metadata that determines how the substream is combined with the main broadcast audio track in the receiving device.

Audio Service	Codec	Substream usage	Channel configuration
Main	AC-3	N/A	Up to 5.1 channels
	Enhanced AC-3	Independent substream 0	Up to 5.1 channels
	aacPlus	N/A	Up to 5.1 channels
Associated			
Delivered in same bitstream as Main audio service	Enhanced AC-3 (main audio must be AC-3 or Enhanced-AC-3)	Independent substream 1 Independent substream 2 Independent substream 3	Up to 5.1 channels
Delivered in separate MPEG-PES	Enhanced AC-3 (main audio must be AC-3 or Enhanced AC-3)	Independent substream 0	Up to 5.1 channels
	aacPlus (main audio must be aacPlus)	N/A	

Table 1: Input formats for a proposed multichannel mixing device

Table 1 shows input formats to such a mixing engine that represent the predominant multichannel audio formats in current and next-generation broadcast systems.

This DMC design supports mixing one main audio service with one associated audio service. The associated service can be delivered in the same Enhanced AC-3 bitstream as the main service through the use of an additional independent substream, or as an Enhanced AC-3 bitstream carried in a separate MPEG PES stream.

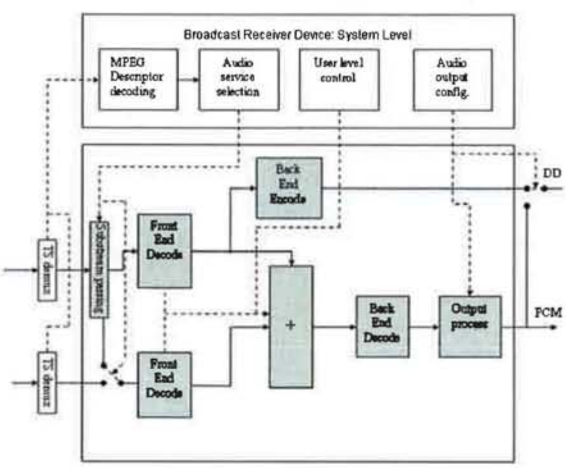

Figure 3: The structure of the decoder-mixer-converter (DMC)

In the case of aacPlus, the associated audio service can be delivered as an aacPlus bitstream carried in a separate MPEG PES stream. Information on the type of associated audio service is carried in the Enhanced AC-3 or aacPlus Descriptor that is found in the program map table of the incoming MPEG-2 Transport Stream.

Each Enhanced AC-3 or aacPlus bitstream used for the broadcast service has an identifying descriptor which the broadcast receiver device uses to inform the user which associated audio services are available. When the user selects a desired service, the DMC in the broadcast receiver device demultiplexes it from the MPEG transport stream.

Example 1

The user selects an associated audio service being carried in a separate MPEG PES associated with the main audio service through its Enhanced AC-3 descriptor. The broadcast receiver demultiplexes both Enhanced AC-3 streams from the broadcast transport stream, and passes them to both inputs of the DMC. The DMC receives information on the user's stream selection from the system layer of the broadcast receiver device and activates both inputs, then decodes and mixes the two streams. The mixing metadata is carried in the Enhanced AC-3 bitstreams containing the associated audio track.

Example 2

The user selects an associated audio service that is being carried in an additional independent substream in the Enhanced AC-3 main audio bitstream. The broadcast receiver demultiplexes only the main audio bitstream, while the DMC receives information on the user's associated service selection from the system layer of the broadcast receiver device and activates its main audio input only. The DMC then uses the substream parser in the main audio chain to separate independent substream 0 (main audio) from the independent substream carrying the associated audio. The associated audio service is then routed inside the DMC to the associated audio input of the mixer, and both independent substreams are decoded and mixed. The mixing metadata is carried in the Enhanced AC-3 bitstreams containing the associated audio soundtrack.

Example 3

The user selects an associated audio service that is being carried in a separate MPEG PES associated with the main audio service through its aacPlus descriptor. The broadcast receiver demultiplexes both aacPlus streams from the broadcast transport stream, and passes these streams to both inputs of the DMC. The DMC receives information on the user's associated service selection from the system layer of the broadcast receiver device, activates both inputs, then decodes and mixes the two aacPlus bitstreams. The mixing metadata is carried in the PES header information of the aacPlus bitstreams and has to be passed to the DMC from the system level parser of the broadcast receiver device.

4.1. Multichannel mixing and the inclusion of system sound effects in surround

As mentioned earlier in this paper, broadcast receivers can go beyond the minimum requirements set forth by DVB to support services for the visually impaired. Implementing a design as shown in Figure X, opens up even more applications for the broadcaster, including the bandwidth-efficient transmission of multiple 5.1 mixes of a program, or enabling a sports event mixed in surround to be enjoyed in different languages or with different commentators for different geographical regions.

A frustration for developers of next-generation DVR devices is that system sounds stored or created in these devices so far can make it out of the box only through their two-channel analog outputs. Typically, the multichannel bitstreams transmitted as part of a program are being passed to the digital outputs (HDMI or S/PDIF) unaltered. As a result, it has been impossible to include system sounds for those outputs. An even more extended design option of the DMC offers inputs for pre-produced system sounds in 5.1 multichannel surround sound. Figure 4 shows the system diagram of the DMC supporting this feature.

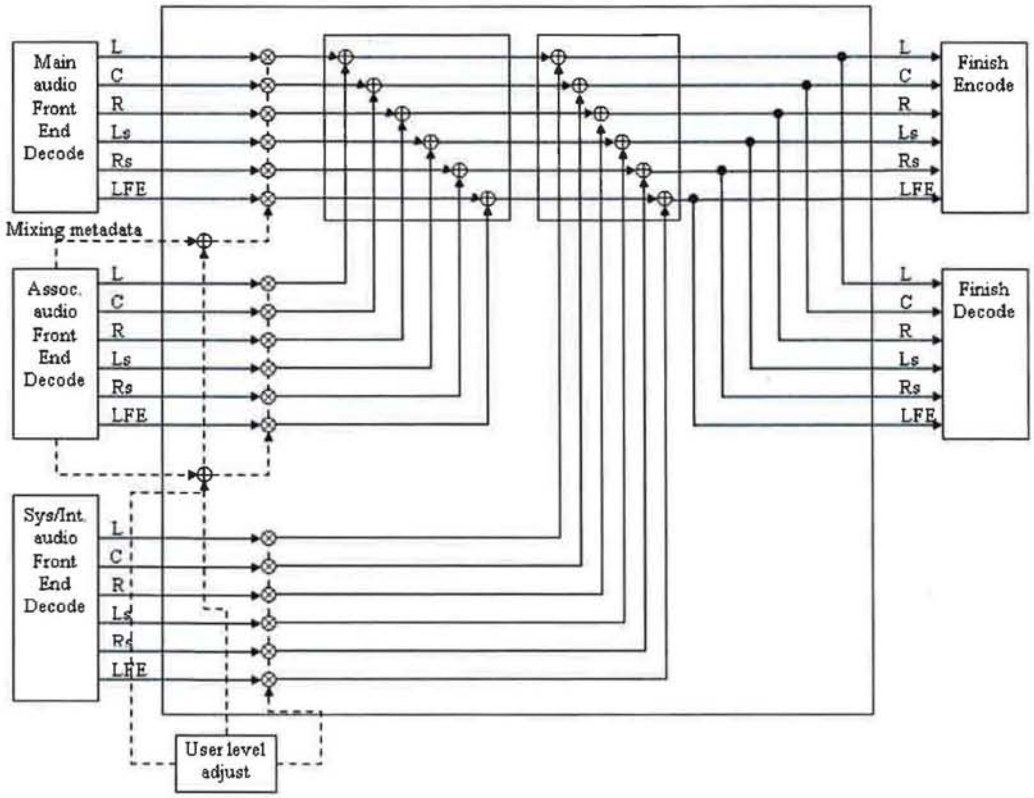

Figure 4: DMC with inputs for multichannel system sounds

This configuration supports mixing of one main audio service with one associated service, plus an additional input for system/interactive sounds which are be stored locally in the set-top box device.

The associated audio service can be delivered in the same Enhanced AC-3 bitstream as the main audio through the use of an additional independent substream, or can be delivered as an Enhanced AC-3 bitstream carried in a separate MPEG PES stream.

Information on the type of associated audio service being delivered is carried in the Enhanced AC-3 Descriptor that is found in the program map table of the incoming MPEG-2 Transport Stream. Each Enhanced AC-3 bitstream present in the broadcast service has an Enhanced AC-3 descriptor associated with it.

The broadcast receiver device uses the contents of each descriptor to inform the user of which associated audio services are available. The user selects the Enhanced AC-3 bitstream or substream carrying the desired associated service. In the case of aacPlus, both Main and Associated Audio services must be carried separate MPEG PES bitstreams.

System/interactive sound samples can either shipped with the broadcast devices, or in the case of a DVR, pushed to the device as data files contained in the MPEG transmission transport stream. This provides a very efficient way to supply high-end customers, who have multichannel home cinema systems and don't use two-channel analog outputs, with the same or even more exciting system sounds that can be mixed-in with the 5.1 broadcast audio.

5. PREREQUISITES FOR THIS NEW MIXING ARCHITECTURE

Support for multichannel mixing and enhanced features such as system sounds in surround call for significant changes in the audio decoder structure of broadcast receiver devices. Many next-generation audio codecs have been designed with these new requirements in mind. One of the two codecs discussed in this paper is the Enhanced AC-3 codec as defined in [2] and [5].

As a successor to the well established AC-3 codec used in many entertainment applications worldwide, Enhanced AC-3 provides a framework for new features like audio description. The following chapter illustrates how Enhanced AC-3 opens opportunities for broadcast operators to new and exciting services while maintaining maximum compatibility with the existing installed base of home cinema systems.

This chapter discusses the requirements for an audio coding scheme to address both short- and long-term technology challenges of next-generation broadcasting such as high-definition television (HDTV). A number of factors have been considered in determining the characteristics of a suitable audio coding scheme:

• Requirements for new broadcast services such as HDTV and services using advanced video codecs such as H.264 or VC-1.

• Opportunities for improving the audio performance of current broadcast services.

• Impact of a new audio coding scheme on broadcast production practices.

• Impact of a new audio coding scheme on consumer hardware and listening environments.

• Differentiation between high-end next-generation and competitive TV services.

6. REQUIREMENTS FOR EXISTING SERVICES, SHORT- AND LONG-TERM

The audio codecs previously specified in [2] offer solutions for many of the requirements facing DTV and IPTV operators—high-quality audio at low bit rates, multichannel audio services, guaranteed connectivity with consumer products through existing IEC 61937 interfaces (S/PDIF or HDMI), and vast improvements to the consumer listening experience through the use of audio metadata. Each codec previously specified in [2] meets some of these requirements, but until recently, there was no single codec that meets all of them.

For example, the AC-3 (Dolby® Digital) codec delivers up to 5.1 channels of audio, offers broadcasters and operators full control over the listening experience for all consumer environments through comprehensive metadata control, and offers standardized connectivity via IEC 61937 to over 40 million existing consumer A/V systems. However, the AC-3 codec is not optimized for low bit-rate performance.

When considering the feature set used by current broadcast services, a new audio codec should offer at least the following:

• Support for mono to 5.1- and 7.1-channel capability.

• Comprehensive metadata support, mandated in both encoder and decoder, with all parameters under encoder control:

- Dialogue normalization to ensure consistent listening levels between programs.
- Downmix capability to ensure backward compatibility with matrix surround, stereo, and mono playback systems.
- Control of dynamic range to ensure optimal reproduction for all consumer listening environments.

• Delivery of discrete 5.1-channel audio to current installed base of A/V receivers via IEC 61937 interfaces, and support for other emerging digital interface standards.

• Improved bit-rate efficiency compared with audio codecs currently in use in DVB and IPTV services, complementing efficiency gains of new video codecs.

• Licensing costs and terms in line with existing audio codecs.

• Encoder and decoder products subject to interoperability testing to ensure consistent performance.

In addition to these requirements for core broadcast services, an opportunity also exists to improve the current provisions for deployment of audio description services for visually and hearing impaired. While relative levels between audio description (AD) and main program services can be controlled both by the broadcaster and the listener, variations in loudness and dynamic range between programs leads to a need for regular adjustments to listening levels by the consumer.

A new audio codec should meet the following requirements to deliver improved AD services:

• Metadata control of dialogue levels to ensure a consistent relative level between main and AD programs.

• Metadata control of the dynamic range of the main program to ensure that AD services are clearly audible at all times.

• Metadata to control mixing of main program and AD services in a broadcast receiver should be supported, to remove the need for frequent manual adjustment of levels in the broadcast receiver. Support for mixing AD services with multichannel as well as stereo program content should also be available.

• Ability to deliver both main program and AD services as a single stream that can be decoded and mixed with a single decoder in the broadcast receiver to simplify implementation in a broadcast receiver.

7. REQUIREMENTS FOR NEW SERVICES

Standards and technologies are being developed for IPTV services and the next generation of DTV broadcasts. New audio codecs must adapt to the requirements of these new technologies. Applications such as high-definition television and interactive services present new opportunities for audio services. A new audio codec should satisfy at least the following requirements to meet the demands of future broadcast services:

• Ability to deliver audio quality improvements to match video quality improvements of high-definition IPTV broadcasting.

• Flexibility to deliver more than 5.1 channels of audio to match future motion picture mixing formats.

• Support for mixing interactive audio content with main program audio, including multichannel audio content.

• Deployment of multiple programs in a single stream, enabling multiple languages, director's commentaries or other advanced services, all controlled by mixing metadata, to be decoded using a single decoder in the broadcast receiver.

7.1. Impact on Broadcast Production Environment

The adoption of a new audio codec for final broadcast should have minimal affect on a broadcaster's working methods, and should not adversely affect or complicate them. In the case of stereo audio services, the process of creating audio content does not differ greatly from codec to codec—the selection of encoding settings need be done only once, based upon the target quality of a broadcaster's service and the capability of the codec selected. If the selected codec supports control of program loudness and dynamics through metadata, this should always be factored into the production process.

When considering multichannel audio services, the workflow of a new emission codec through the broadcast production environment must be carefully considered. Multichannel audio presents a number of challenges to a broadcaster. These challenges include distribution of 5.1- (or higher) channel audio content within a broadcast facility equipped only for stereo content, and the task of creating audio metadata to ensure that 5.1 programming can deliver optimal backward compatibility with all listening environments.

When considering the integration of a new audio codec with a broadcast production environment, it is important that the level of metadata functionality should at least match and preferably exceed that of current codecs, to maintain a broadcaster's ability to deliver consistent audio quality. A simple interface between multichannel production equipment and transmission encoders for both audio and metadata is desirable.

Today, the Dolby E format is predominantly used in multichannel audio production. This format can handle up to eight channels of audio in a single AES/EBU carrier and integrates the complete metadata functionality as described in [1]. As a consequence, most contribution of 5.1 or 7.1 surround sound audio in broadcast production is done in Dolby E. IPTV operators will be able to benefit from its flexibility for both linear turn-around of existing broadcast channels as well during the preparation of assets for non-linear offerings such as video-on-demand.

In order to support audio description services, Dolby E can be used to store both the main audio soundtrack (up to 5.1 channels) along with an audio description track (up to 2 channels) within the same Dolby E bitstream. This bitstream can be stored on any carrier offering transparent AES/EBU audio capabilities at 20bit resolution.

7.2. Impact on Consumer Products and Listening Environments

A new audio codec should offer performance improvements for the consumer, while ensuring simple integration into the current consumer listening environment, and offer flexibility for future developments in consumer product enhancements. To ensure a consistent listening experience for all consumers, a new audio codec should meet the following requirements:

• Decoders must maintain compatibility with existing consumer A/V receivers and IEC 61937 interfaces when delivering discrete 5.1-channel content, without introducing excessive complexity to the decoder design.

• All decoders must be able to receive and decode multichannel audio services to deliver either a matrix surround, stereo, or mono downmix as required, removing the need for audio simulcasting (the new audio codec should not remove the possibility of audio simulcasting, only the need).

• Decoder complexity should be in line with current designs for an equivalent feature set.

• Metadata created during the production process must be supported by all decoders. If it is not, the benefits of using metadata are usually lost.

• The new codec should be compatible with emerging and future digital interface standards without introducing excessive complexity to the decoder design.

• Licensing costs and terms in line with existing audio codecs.

8. MEETING THESE REQUIREMENTS

In consideration of these requirements, Dolby developed Enhanced AC-3 (Dolby Digital Plus) for use in next-generation applications. This coding scheme has already been included in [5] and has been selected as mandatory technology for HD DVD players and as optional technology for Blu-ray Disc players. The scheme has also been standardized by DVB for next- generation broadcast and IPTV services in [2] and [4][4].

Enhanced AC-3 is well suited in applications that include lower data-rate carriage of audio and its conversion to the AC-3 coding standard for playback on today's installed base of audio/video entertainment equipment. It also supports combining streamed content with a main audio program in interactive multimedia; the reproduction of greater than 5.1 channels for playback of both existing and future cinema content; and the efficient transcoding of AC-3 program content to lower data-rate Enhanced AC-3 bitstreams, and conversion back to AC-3 for playback on the very large installed base of consumer Dolby Digital decoders.

8.1. Compatible Lower Data-Rate Carriage

Growing number of applications not only require lower data rates, but also compatibility with the existing broadcast-reception and audio/video decoding infrastructure.

The Enhanced AC-3 system is an excellent solution for these applications because of its inherent lower tandem coding losses compared to AC-3 and its greater coding efficiency provided by new coding tools. This results in minimizing quality losses when large content libraries are transcoded to other formats for IPTV or other advanced service environments. To ensure compatibility with the large installed base of Dolby Digital decoders, Enhanced AC-3 enables low-loss conversion to standard AC-3 over a digital audio interconnect such as S/PDIF and decoding by a standard Dolby Digital decoder.

The conversion stage is a special form of transcoder that minimizes quality degradations resulting from tandem coding losses, which is possible by using the same filterbank, transform block alignment, bit-allocation process, and basic framing structure as conventional AC-3.

8.2. A Next-Generation Broadcast Receiver Device

The next-generation broadcast receiver application is very similar to the conventional AC-3 reception paradigm, except the need for greater video channel capacity for HDTV services requires the transmission of audio programming at lower data rates than AC-3 applications. Traditionally, AC-3 has been deployed at 192–256 kbps for stereo and 384–448 kbps for 5.1-audio applications. The use of the new coding tools in Enhanced AC-3 allows for lower data rates while permitting efficient conversion to a conventional AC-

3 bitstream at 640 kbps for tested compatibility with existing home theaters. Figures 1 through 4 show the different configurations and use cases of this converter/decoder.

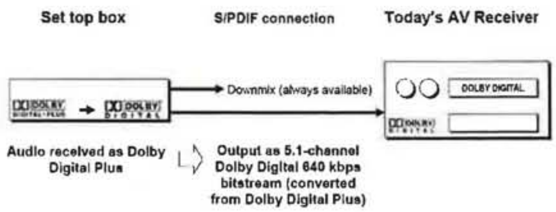

Figure 5: Scenario 1: Broadcast receiver device decoder-converter: Enhanced AC-3 input signal, S/PDIF connection (mixed or unmixed)

In scenario 1, the Enhanced AC-3 decoder-converter converts an incoming Enhanced AC-3 signal into a standard AC-3 signal at 640 kbps for output over IEC61937 interfaces. This conversion maintains all metadata information within the Enhanced AC-3 bitstreams and passes it on to the AC-3 bitstream. Channel configurations between mono and 5.1 remain unchanged. If the incoming audio had a 7.1-channel configuration, the decoder-converter would extract the 5.1-channel core from the signal and pass it on as AC-3 in a 5.1-channel configuration at 640 kbps.

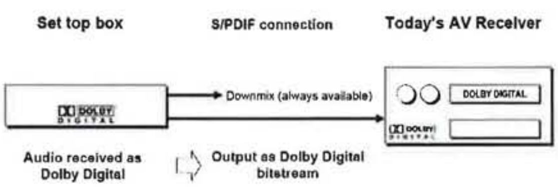

Figure 6: Scenario 2: Broadcast receiver device decoder-converter: AC-3 input signal, S/PDIF connection (mixed or unmixed)

In scenario 2, the Enhanced AC-3 decoder-converter acts as a standard AC-3 decoder and provides a two-channel downmix from the incoming audio while simultaneously passing the incoming stream to the IEC61937 interfaces unchanged as the AC-3 multichannel decoding capability can always be assumed in any digital A/V receiver.

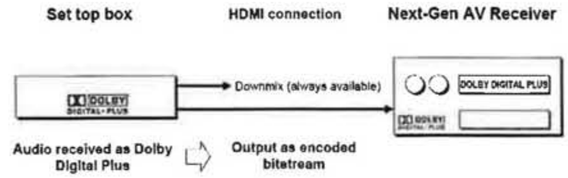

Figure 7: Scenario 3: Broadcast receiver device decoder-converter: Enhanced AC-3 input signal, HDMI connection (mixed or unmixed)

Scenario 3 is only possible in a situation where the multichannel home theater device supports internal Enhanced AC-3 decoding. In this case, no conversion is necessary and the set-top box passes the Enhanced AC-3 stream unchanged while simultaneously creating a two-channel downmix for its stereo interfaces. This scenario enables the use and playback of 7.1-channel configurations. A/V receivers that feature HDMI input without internal Enhanced AC-3 decoding capability will automatically communicate this to the set-top box device via the negotiation protocol on the HDMI link. Hence, the Enhanced AC-3 decoder-converter will automatically provide a standard AC-3 bitstream on the HDMI interface.

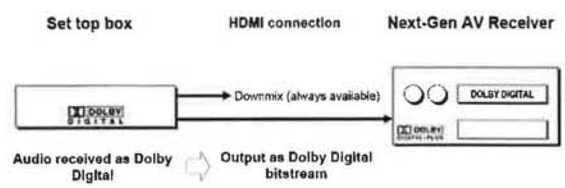

Figure 8: Scenario 4: Broadcast receiver device decoder-converter: AC-3 input signal, HDMI connection (mixed or unmixed)

Scenario 4 replicates scenario 2 with the exception of an HDMI interface connection between the set-top box device and the A/V receiver. There is a standard AC-3 input signal in both cases.

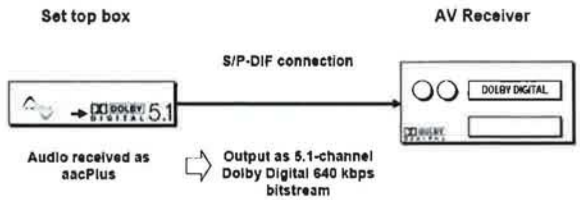

Figure 9: Scenario 5: Broadcast receiver device decoder-converter: aacPlus input signal, S/PDIF connection (mixed or unmixed)

The devices shown in Figure 5 through Figure 9 accept both the Enhanced AC-3 and AC-3 bitstreams and always output the appropriate version of AC-3 for the respective A/V receivers as the consumer switches programs within the TV service. The greater efficiency of the Enhanced AC-3 system allows for a greater number of programs within the broadcast system, while preserving full functionality for legacy receiver hardware and enabling new and extended functionality for new receiver hardware.

9. CONCLUSION

Audio description services are becoming a standard feature on next-generation broadcast platforms. Some European countries even mandate AD services on an increasing number of programs offered by broadcasters, operators and programmers. This paper shows how the support for these services through an enhanced mixing module integrated into a multichannel decoder for next-generation broadcast codecs used in broadcast receivers offers features that go beyond the mandate. These new features offer full home cinema compatibility and therefore enhance the entertainment experience for all audiences.

At the same time, 5.1 surround sound is an important component of the HDTV experience. Existing MPEG2-based HDTV services in the U.S., Europe, and Australia already offer 5.1 surround sound using Dolby Digital technology as documented by DVB and ATSC. With the move to next-generation video coding systems, broadcasters and operators also have many new requirements for audio delivery. This paper shows how new features of the Enhanced AC-3 aacPlus systems meet these requirements while also maintaining compatibility with the more than 40 million existing consumer home cinema systems.

10. REFERENCES

[1] Standards document ETSI TS 102 366, V1.1.1, "Digital Audio Compression (AC-3, Enhanced AC-3) Standard," February 2005.

[2] Standards document ETSI TS 101 154, V1.8.1, "DVB: Implementation Guidelines for the Use of Video and Audio Coding in Broadcasting Applications," July 2007.

[3] Louis D. Fielder et al., "Introduction to Dolby Digital Plus, an Enhancement to the Dolby Digital Coding System," 117th AES Convention, October 2004.

[4] ETSI TS 102 005 V1.3.1 "DVB: Specification for the Use of Video and Audio Coding in DVB Services Delivered Directly over IP Protocols".

[5] Advanced Television Systems Committee: Digital Audio Compression Standard (AC-3, E-AC-3) Revision B; Document A/52B, 14, June 2005.

[6] ITU-T Recommendation H.222.0 / ISO/IEC 13818-1: "Information Technology - Generic Coding of moving pictures and associated audio information: Systems"- and ITU-T Recommendation H.222.0 / Amendment 3 (2004)/ISO/IEC 13818-1: 2000/Amendment 3:(2004): "Transport of AVC video data over ITU-T Recommendation H.222.0 /ISO/IEC 13818-1 streams".

Appendix B

European HD Specifications:

	Receiver Mixed		
	O: optional M: mandated	E-AC-3	Method
UK D-Book v 6.2.1(*)	M (22.3.1.4)	M (4.5.5)	Single PID and Dual PID
Italy HD-Book v2.0	M (6.1.1.1)	M (6.1.1.1)	Single PID and Dual PID
France HDForum v2.5e	O (4.7)	M (4.7)	O Single PID and Dual PID
Ireland v2.0	O (5.3.7)	O (5.3.7)	Single PID
Nordig v2.2(*)	O (6.2.5) Annex F		
EBU Tech 3333(*)	O (5)	O (5)	

(*) indicates updates are in progress
O indicates Optional
M indicates Mandatory